April 16, 2002 LBNL-49602

A Bell-type Theorem Without Hidden Variables *

Henry P. Stapp
Lawrence Berkeley National Laboratory
University of California
Berkeley, California 94720

Abstract

Bell's theorem rules out local hidden-variable theories. The locality condition is the demand that what an experimenter freely chooses to measure in one space-time region has no influence in a second space-time region that is spacelike separated from the first. The hidden-variable stipulation means that this demand is implemented through requirements on an assumed-to-exist substructure involving hidden variables. The question thus arises whether the locality condition itself fails, or only its implementation by means of the assumed hidden-variable structure. This paper shows that any theory that satisfies two generally accepted features of orthodox quantum theory and that yields certain predictions of quantum theory cannot satisfy the aforementioned locality condition. These two features are that the choices made by the experimenters can be treated as localized free variables and that such free choices do not affect outcomes that have already occurred.

*This work was supported in part by the Director, Office of Science, Office of High Energy and Nuclear Physics, of the U.S. Department of Energy under Contract DE-AC03-76SF00098.

1. Introduction.

The premises of Bell's original theorem[1] postulate the existence of a substructure that determines in a local manner the outcomes of the whole set of alternative possible measurements at most one of which can actually be performed. The implementation of the locality condition in this way thus involves technical "hidden-variable" assumptions that go beyond the locality condition itself. Consequently, Bell's proof of the inconsistency of this local hidden-variable assumption with certain predictions of quantum theory casts no serious doubt on the locality condition itslf: its implementation via the hidden-variable structure would appear to be the more likely cause of the inconsistency.

Bell[2] introduced later a seemingly weaker local hidden-variable assumption, but this latter form can be shown[3,4] to entail the original one, apart from errors that tend to zero as the number of experiments tends to infinity. Thus both forms of the assumption place on the class of theories covered strong mathematical conditions that are essentially equivalent to the assumption that values can be pre-assigned conjunctively and locally to all of the outcomes of all of the alternative possible measurements. That assumption conflicts with orthodox quantum thinking. Thus these hidden-variable theorems place in no serious jeopardy the primitive locality condition that a free choice made by an experimenter in one space-time region has no influence in a second region that is space-like separated from the first.

The present paper shows that this locality condition fails in a much larger class of theories, namely those that are compatible with the properties of Free Choice and No Backward-in-Time Influence on observed outcomes, and that yield certain predictions of quantum theory in experiments of the Hardy type[5]. The first two of these three properties are now described.
Free Choices.
For the purposes of understanding and applying quantum theory, the choice of which experiment is to be performed in a certain space-time region can be treated as an independent free variable localized in that region. Bohr repeatedly stressed the freedom of the experimenter to chose between alternative possible options. This availability of options is closely connected to

his "complementarity" idea that the quantum state contains complementary kinds of information pertaining to the various alternative mutually exclusive experiments that might be chosen. Of course, no two mutually incompatible measurements can both be performed, and an outcome of an experiment can be specified only under the condition that that particular experiment be performed.

No Backward-in-Time Influence. (NBITI) *An outcome that has already been observed and recorded in some spacetime region at an earlier time can be considered fixed and settled, independently of which experiment a far-away experimenter will freely choose to perform at some later time.* This assumption assigns no value to a local measurement except under the condition that this local measurement be performed. But any such locally observed value is asserted to be independent of which measurement will at some later time be freely chosen and performed in a spacelike separated regions.

This NBITI assumption is required to hold in at least one Lorentz frame of reference, hence forth called LF.

This NBITI assumption is compatible with relativistic quantum field theory. In the Tomonaga-Schwinger[6, 7] formulation the evolving state is defined on a forward moving space-like surface. Their work shows that this surface can be defined in a continuum of ways without altering the predictions of the theory, so that no Lorentz frame is singled out as preferred. On the other hand, their formalism *allows* the quantum state to be defined by the constant time surfaces in any one single Lorentz frame that one wishes to choose, and shows that in this one frame the evolution, including all reductions associated with specific outcomes of measurements, proceeds forward in time, with a well defined past that is not influenced either by later free choices made by experimenters or by the outcomes of the later measurements.

This NBITI assumption is a small part of the larger locality condition in question here, which is the demand that what an experimenter freely chooses to do in one region has no effect in a second region that is spacelike separated from the first.

These definitions allow the following theorem to be stated:

Theorem. Suppose a theory or model is compatible with the three premises:

1. Free Choices, which assert that the choice made in each region as to which experiment will be performed in that region can be treated as a localized free variable,
2. No Backward in Time Influence, which asserts that experimental outcomes that have already occurred in an earlier region can be considered to be fixed and settled independently of which experiment will be chosen and performed later in a region spacelike separated from the first, and
3. For each of the alternative possible combinations of the free choices of which measurements will be performed in a Hardy-type experiment the predictions of quantum theory will hold under the condition that that combination is chosen.

Then this theory or model violates the following Locality Condition:
The free choice made in one region as to which measurement will be performed there has, within the theory, no influence in a second region that is spacelike separated from the first.

2. Proof of the Theorem.

The theorem refers to the following Hardy-type [5] experimental set-up.

There are two experimental spacetime regions R and L, which are spacelike separated, with L lying earlier than R in LF. In region R there are two alternative possible measurements, $R1$ and $R2$. In region L there are two alternative possible measurements, $L1$ and $L2$. Each local experiment has two alternative possible outcomes, labelled by $+$ and $-$.

The detectors are assumed to be 100% efficient, so that if a measurement is chosen in R then some outcome of that measurement, either $+$ or $-$, will appear in R, and if a measurement is chosen in L then some outcome of that measurement, either $+$ or $-$, will appear in L.

The three assumptions are implemented by asserting that:
1. For each of the two choices $L1$ or $L2$ available to the experimenter in L, and for each of the two alternative possible outcomes $+$ or $-$ of that experiment, there are *instances* in which that experiment, $L1$ or $L2$, is performed, that

outcome, $+$ or $-$, of that experiment appears, and an associated pair of options $R1$ and $R2$ exists, and
2. In each such instance the predictions of quantum theory hold if $R1$ is performed, and the predictions of quantum theory hold if $R2$ is performed.

These conditions can be formalized by defining an *instance* i to be a triad $\{X_i, W_i(R1), W_i(R2)\}$, where X_i is a possible outcome of a possible measurement in L, and $W_i(R1)$ is a possible world in which X_i appears in L and $R1$ is performed in R, and similarly for $R2$, and then specifying that the predictions of quantum theory hold in $W_i(R1)$ and in $W_i(R2)$. This structure just expresses in a compact notation the ideas that there is a free choice and for each choice two possible outcomes in region L, and that for any possible outcome in L there are two options, R1 and R2, available to the experimenter in R, and that we are considering a theoretical structure that allows one to say, for each of these two options in R, that if that option is chosen then the corresponding predictions of quantum theory about outcomes in R will hold. These assumptions are built into hidden-variable theories that reproduce the predictions of quantum theory, but the local hidden-variable theories have additional structure that is not entailed by the concept of an *instance*, which is compatible both with normal quantum philosophy, and with all the predictions of quantum theory.

The first two pertinent predictions of QT for this Hardy setup are these(*):

For every instance i such that X_i is a possible outcome of performing $L1$,

(2.1): If X_i is $-$ then outcome $-$ appears in R in $W_i(R1)$.
(2.2): If X_i is $+$ then outcome $-$ appears in R in $W_i(R2)$.
These two conditions immediately entail:

Property 1. In every instance i such that X_i is a possible outcome of performing $L1$:
Outcome $-$ appears in R in $W_i(R1)$,
OR
Outcome $-$ appears in R in $W_i(R2)$.

Proof of Property 1. For every instance is such that X_i is an outcome of performing $L1$, as specified in Property 1, either $X_i = -$ or $X_i = +$. For instances i of the first kind prediction (2,1) ensures that the first alternative specified in Property 1 is satisfied. For instances i of the second kind prediction (2.2) ensures that the second alternative is satisfied. Q.E.D.

The second two pertinent predictions of QT for this Hardy setup are:

For every instance i such that X_i is a possible outcome of performing $L2$,

(2.3): If X_i is $-$ then outcome $+$ appears in R in $W_i(R1)$.

There are instances i such that X_i is a possible outcome of performing $L2$ and

(2.4): $X_i = -$ and the outcome $+$ appears in R in $W_i(R2)$.

These two conditions immediately entail:

Property 2. There are instances i such that X_i is a possible outcome of performing $L2$, and
Outcome $+$ appears in R in $W_i(R1)$,
AND
Outcome $+$ appears in R in $W_i(R2)$.

Proof of Property 2. Prediction (2.4) entails that there are instances i such that the second of the two conditions demanded by Property 2 is satisfied. Prediction (2.3) then ensures that for any of those instances i the first of the two conditions demanded by Property 2 is also satisfied. Q.E.D.

Properties 1 and 2 are incompatible with the locality condition that the choice between $L1$ and $L2$ has no influence in R. For under the condition that $L2$ is performed there are instances i that satisfy the condition on possible outcomes in region R that is specified in Property 2, whereas if $L1$ is performed then the condition specified in Property 1 must hold. But these two conditions on outcomes in R are converse conditions. This condition that logically *converse conditions* on the outcomes in R hold according to whether $L1$ or $L2$ is freely chosen in L is a sufficient condition for the existence within the theory of an influence from L to R.

3. The Concept of Influence

There are perhaps many possible meanings of the concept of "influence", and hence it is important to be clear about the one being used here.

I am adhering here to what I believe to be the "orthodox quantum philosophy." What I mean by this is that one must make no assumption that is mathematically or logically equivalent to the assumption that the outcome of an unperformed local measurement operation exists, or to the assumption that an essentially classical-type reality exists. It also means that our physical theories are to be regarded as essentially rules that we use to make predictions about our future experiences on the basis of the information that we could get by performing possible measurements now. Thus the subject matter of both theoretical physics and the present study is the logical and mathematical structure of such theories, not speculations about some fundamentally unknowable ontological structure that lies behind the empirical phenomena.

The purpose of a nonlocality theorem is to exhibit structural properties of theories that reproduce the predictions of quantum theory and that conform to certain other conditions about the structure of the theory. The first of these other condition imposed here is the idea that choice of which of the two local measurements will be performed is to be treated as a free variable. I assume, in accordance with orthodox thinking, that in the actual world only one choice is made in each region on any actual occasion. Hence we must go to the realm of theoretically possible worlds even to formulated the idea of nondependence on free choice that we are trying to impose. Of course, theoretical physics has traditionally been recognized as dealing with a whole class of theoretically possible worlds, not merely with the one unique physical world that we actually inhabit.

The essential ingredient of any proof of the nonlocality property (i.e., the faster-than-light-influence or faster-than-light transfer of information property) is the tandem or simultaneous use of *both* of the two free variables:

one can always accommodate the predictions of quantum theory by allowing faster-than-light influences in just one direction. What cannot be done is to demand, in addition to the no backward-in-time-influence assumption, that for *both* of the two later possible experiments, *at most one of which can actually be performed*, the predictions of quantum theory hold. One can never empirically demonstrate the failure of this theoretical assumption because one could always assert that the predictions fail in the experiment that was not performed. Thus the theorem does not assert that nature herself must involve faster-than-light action. It says rather that any *theoretical conception of nature* that reproduces the predictions of quantum theory, and the no-bachward-in-time-influence condition, cannot be required to reproduce the predictions of quantum theory in *both* of the theoretically possible experiments between which the later experimenter is supposedly free to choose, without contradictory conditions holding in the later region under the two alternative possible condition between which the experimenter in the earlier spacelike separated earlier region is free to choose.

I have defined the arising of *contradictory conditions* in the later region under the two alternative choices made in the earlier region to be a sufficient condition for the existence of an *influence* in the later region of the choice made in the earlier region.

The conclusion obtained here about faster-than-light influences parallels Bohr's reply to the paper of Einstein, Podolsky, and Rosen[8]. The assumption of those authors was that there was no faster-than-light influence of any kind. Bohr's response[9] was a partial denial of that assumption: he granted that there was no faster-than light "mechanical disturbance" , but noted that "there is an influence on the very conditions that define the possible types of predictions regarding the future behavior of the system." A "mechanical disturbance" would be one capable of transmitting a signal, whereas the others pertain to "predictions", and hence to the *theoretical structure*, which from Bohr's point of view, was primarily a tool for making predictions. Bohr's recognition of the existence of faster-than-light influences of this latter kind was, like the one deduced here, made completely within the framework of the Copenhagen interpretation of quantum theory. In the argument presented here the condition that is influenced from afar is of the kind considered by Einstein, Podolsky, and Rosen: it involves two *alternative possible* future situations, whereas "mechanical disturbances" involve only actually performable sequences of experiments.

4. Conclusions

All the assumptions of this nonlocality theorem are compatible with orthodox quantum philosophy, and the conclusion is compatible with relativistic quantum field theory. The theorem therefore covers orthodox quantum theory as a special case. No logical contradiction occurs: the only conflict is with certain ideas carried over from classical relativity theory.

The locality condition whose violation is demonstrated here is similar to the one occurring in Bell's theorems in that: (1), the dependence in question is on an experimenter's free choice of which measurement to perform in a certain region, and (2), the property that depends on this free choice depends jointly on *both* of the alternative possibilities between which the experimenter in the other region is free to choose. The present approach constructs the structure needed for the proof, namely the notion of instances, directly from the assumptions of Free Choice and No Backward in Time Influence, and then carries through the proof, without generating any contradiction with quantum theory, or using any combination of properties that are logically equivalent to Bell's hidden-variable assumptions. By dispensing in this way with the hidden-variable substructure the present theorem evades challenges to Bell's theorems of the kind that recently appeared in the Proceedings of the National Academy of Science[10].

*[NB: To obtain these four predictions from Hardy's paper, one transforms my notation into Hardy's using

$$(L, R) \to (1, 2)$$

$$(1, 2) \to (U, D)$$

$$(+, -)_L \to (0.1)$$

$$(+.-)_R \to (1, 0)$$

and uses the three zero's connecting my pairs of states $(R1+, L1-)$, $(L2-, R1-)$ and $(L1+, R2+)$ that arise from his Eqs. $(13.a, b, c)$ to obtain my (2.1),(2.2), and (2.3), respectively, and uses his $(13.d)$, which says that my matrix element $(L2-, R2+)$ is positive, to obtain my (2.4).]

Acknowledgement

Extensive discussions with Asher Peres, Philippe Eberhard, Jerry Finkelstein, and particularly Abner Shimony, have contributed importantly to the development of this paper.

References

1. J.S. Bell, Physics, **1**, 195 (1964);
J. Clauser and A. Shimony, Rep. Prog. Phys. **41**, 1881 (1978).
2. J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge Univ. Press, 1987, Ch. 4.
3. H. Stapp, Epistemological Letters, June 1978. (Assoc. F Gonseth, Case Postal 1081, Bienne Switzerland).
4. A Fine, Phys. Rev. Lett. **48**, 291 (1982).
5. L. Hardy, Phys. Rev. Lett. **71**, 1665 (1993); P. Kwiat, P.Eberhard A. Steinberg, and R. Chiao, Phys. Rev. **A49**, 3209 (1994).
6. S. Tomonaga, Prog. Theor. Phys. **1**, 27 (1946).
7. J. Schwinger, Phys. Rev. **82**, 914 (1951).
8. A. Einstein, B. Podolsky, and N. Rosen. Phys. Rev. **47**, 777 (1935).
9. N. Bohr, Phys. Rev. **48**, 696 (1935).
10. K. Hess and W. Philipp, PNAS **98** 14224, 14233 (2001).